DOCUMENT RESUME

ED 046 657 RE 003 277

AUTHOR Rankin, Earl F.
TITLE Grade Level Interpretation of Cloze Readability Scores.
PUB DATE Dec 70
NOTE 12p.; Paper presented at the National Reading Conference, St. Petersburg, Fla., Dec. 3-5, 1970
AVAILABLE FRCM Twentieth Yearbook of the National Reading Conference, Inc., Marquette University, 1217 W. Wisconsin Ave., Milwaukee, Wis. 53233

EDRS PRICE EDRS Price MF-$0.65 HC Not Available from EDRS.
DESCRIPTORS *Cloze Procedure, *Elementary School Students, Informal Reading Inventory, *Readability, Reading Level, Reading Materials, *Reading Research

ABSTRACT

Two studies were conducted to investigate the use of a technique for providing a grade-level interpretation for cloze readability scores. In the first study, Bormuth's cloze-test criteria were compared with the results of the Fry and Dale-Chall readability formulae. A total of 133 students in grades 4 to 8 were required to read a 250-word passage with every fifth word deleted. The results indicated a fairly close correspondence between Bormuth's 44 percent criterion for cloze readability analysis and the readability formulae. In the second study, a method devised by the writer was used. The Informal Cloze Readability Inventory (ICRI) uses materials of unknown difficulty to be read by readers of known reading ability in order to evaluate the readability of a book. The ICRI and the Paragraph Meaning subtest of the Stanford Reading Achievement Test were given to groups of six average readers in grades 4 to 6. When cloze instructional levels were analyzed, they did not appear to be very different from the readability levels as measured by the readability formulae. It was concluded that valid readability assessments can be made using the ICRI based on a small number of subjects. References are included. (DH)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

Grade Level Interpretation of Cloze Readability Scores*

Earl F. Rankin
University of Kentucky

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL BY MICROFICHE ONLY HAS BEEN GRANTED BY

Earl F. Rankin

TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

ED046657

The purpose of this study was to provide a grade level interpretation for cloze readability scores using a simple technique that can be carried out by classroom teachers. The objective was to study practical methods by which a teacher could use the cloze procedure in evaluating the readability of a passage (or book) and assign to this passage a minimum reading grade level which a student in her class must achieve in order to comprehend the passage. If this can be accomplished without too great an expenditure of time and effort, the use of the cloze procedure for this purpose would have much to recommend it over the use of conventional readability formulae.

In contrast to readability formulae, the cloze procedure measures readability directly from a person, not from material alone. It measures specific contemporary target groups, and thus it need not rely on norms established in the past. It measures readability in relation to a given background of experience and in a motivational context. Unlike many formulae, it is not fooled by long, easy sentences or short, hard words. It assesses language correspondence between the author and the target group. It reflects the redundency in the passage. In so far as factors like organization and subtle elements of style influence comprehension, this is reflected in a cloze score. It has even been found by Rankin and Culhane (1970) that pictures may influence cloze scores. Furthermore, cloze tests are easily constructed and scored, and they do not require complicated statistical manipulations in order to be interpreted.

RE003 277

*The author wishes to thank graduate students Pamela Murphy and Joanne DeWitt for their invaluable assistance in gathering and analyzing the data for this paper.

Many studies--Taylor (1953), Sukeyori (1957), Bormuth (1962), Gallant (1964), etc.--have shown that the cloze procedure is a valid measure of readability. Indeed, it has been used by Bormuth (1969b) as the validity criterion for evaluating other readability measurements. For sometime the main weakness of the cloze procedure as a measure of readability was the absence of criteria for interpreting raw scores. The relative difficulty of two or more passages could be determined, but no interpretation could be placed upon the difficulty of each passage. However, Bormuth (1967, 1968) established criterion reference scores for interpreting cloze test scores in relation to conventional multiple-choice comprehension test scores. His 1968 results, which he considers to be the most valid, showed that a cloze score of 44% was comparable to a 75% comprehension score (i.e. instructional level) and a cloze score of 57% was comparable to a comprehension score of 90% (i.e. independent level). Rankin and Culhane (1969) replicated this work and obtained essentially the same results using different materials. Thus, cloze test scores may now be interpreted in relation to these conventional reference points in selecting materials for use in the classroom.

It should be noted, however, that these criteria are quite arbitrary. There is no real justification for using these particular percentages as criteria for selection. In another important study, Bormuth (1969a) sought to determine an empirical basis for establishing instructional level. Using "information gain" as a criterion, he found that subjects were able to achieve a maximum amount of information gained by reading materials with a cloze score of 37% (i.e. 65% multiple-choice comprehension). Little more gain was achieved by reading easier materials. He cautioned against ready acceptance of this readability criterion since this study was a preliminary step in the direction of establishing an empirically based criterion for

readability. He also observed that students experienced some degree of frustration reading materials at this level. In the present study, all three of Bormuth's criteria will be investigated as standards for comprehension of comprehensibility of materials.

Still lacking in measuring cloze readability is a method to assign reading grade level interpretations to cloze scores. A research technique using regression equations to predict standardized reading grade scores from cloze scores has been used by Bormuth (1969b) to assign grade placement scores to passages. However, this has no utility for the practioner. The present investigation has studied the results of a "rough and ready" method of assigning grade level scores to passages by relating the grade level reading test scores of the pupils to one or more of Bormuth's criterion reference scores. First, a study was done using an entire class. Second, a preliminary investigation was carried out using a small group of pupils selected on the basis of reading test grade scores.

Whole Class Procedure

The purpose of this procedure was to determine the grade level readability for several passages by administering cloze tests based upon the passages to an entire class. The study sought to find out what type of results would be obtained by using Bormuth's cloze criterion reference scores and reading grade levels based upon standardized reading tests to determine the grade level readability of several passages. The results of this procedure were compared with the results of several readability formulae.

Subjects: Five classes of public school students in successive grades from four through eight were used as subjects. The average class size was 26. A total of 133 students were involved in the investigation.

Materials: Five articles about birds were chosen from several books on

this topic. A 250 word passage for cloze test construction was selected from the middle of each article. In constructing the test, the first and last sentences were left intact. Starting with the second sentence, every fifth word was deleted and replaced with an underlined blank space of fifteen spaces. By this procedure, a 50 item cloze test was produced.

Data Gathering Procedures: Each of five classes was given a different cloze test during the middle of the spring semester. Students were given as much time as necessary to complete the test. The length of time was between thirty and forty-five minutes for all classes. Recent grade level scores were obtained for all students on the Paragraph Meaning sub-test of the Standford Reading Achievement Test. Each 250 word passage was evaluated by both the Dale-Chall and Fry readability formulae.

Analysis: Cloze tests were scored by giving credit for only the exact word that had been deleted. Misspelled words were not counted wrong. The reading grade level scores of students were averaged for those students making cloze test percentage scores of 38, 44, and 58.[1]

Results: A comparison of cloze grade level readability scores with those obtained by the Dale-Chall and Fry readability formulae is presented in Table I.

If the 38% criterion (i.e. 65% multiple-choice comprehension) is a valid one, it can be concluded that the readability of these passages is easier as measured by the cloze procedure than as measured by the two formulae. It may be that the 75% comprehension criterion is too stringent. However, more work is needed on the empirical determination of the instructional reading level before we can be sure of the appropriate criterion. In any case, the cloze procedure, unlike readability formulae, can be flexibly adapted to the new criteria as determined by research.

---

[1]The 38% and 58% were rounded off from Bormuth's 37% and 57% criteria because the latter criteria corresponded to fractions of raw scores on this test of 50 items.

Table I

Grade Level Readabilities of Five Passages
as Measured by the Cloze Procedure
and Two Readability Formulae

CLOZE TEST AT 38%					
Grade Level of Subjects	4	5	6	7	8
Cloze Test	4.9	5.5	6.3	5.3	6.8
Fry Formula	6	7	7	5	8
Dale-Chall Formula	5-6	7-8	5-6	7-8	11-12
CLOZE TEST AT 44%					
Grade Level of Subjects	4	5	6	7	8
Cloze Test	5.9	5.6	8.8	6.8	7.6
Fry Formula	6	7	7	5	8
Dale-Chall Formula	5-6	7-8	5-6	7-8	11-12
CLOZE TEST AT 58%					
Grade Level of Subjects	4	5	6	7	8
Cloze Test	*	*	*	10.6	8.7
Fry Formula	6	7	7	5	8
Dale-Chall Formula	5-6	7-8	5-6	7-8	11-12

*No scores were obtained at this percentage on these particular tests.

Using the 44% criterion (i.e. 75% multiple-choice comprehension), it is clear that there is a fairly close correspondence between results of the cloze readability analysis and the two formulae. When discrepancies between cloze and formulae results appear, it is believed (for reasons previously given) that the cloze results are more valid. For example, the greatest discrepancy exists between cloze test results and Dale-Chall results for the eighth grade group (i.e. 7.6 cloze verses 11-12 Dale-Chall). In this case, the Dale-Chall results are probably invalidated by the constant repetition of several words like "migration" and "seasonal" which are not on the Dale list. The cloze results are not "fooled" by this factor.

In using the 58% criterion (i.e. 90% multiple-choice comprehension) we can evaluate the independent reading level by using the cloze procedure. This is not possible using readability formulae. For grades four, five, and six, these passages were too difficult for anyone to obtain the 58% criterion. Therefore, using the cloze procedure, we find that the articles are too difficult for independent reading. Most of the formulae results point to the difficulty of these passages for grades four, five, and six, but do not permit a conclusion about independent level. The seventh grade article is apparently suitable for independent use at grade 10. This conclusion differs considerably from the results of the formulae.

In conclusion, it appears that the cloze readability levels using the 44% criterion are not unlike formulae results in most cases, at least for this sample. In atypical groups (i.e. culturally disadvantaged pupils), cloze results could be expected to be considerably different from formulae results. Assessments of independent level can be made with the cloze procedure but not with conventional formulae. Finally, since both the conventional criteria for instructional and independent reading levels have not been empirically determined,

the cloze procedure can be flexibly adapted to new criteria determined by careful research.

Small Group Procedure

Another method involving a small number of subjects to determine grade level readability with the cloze procedure was devised by the writer. This method is called the Informal Cloze Readability Inventory (I.C.R.I.). This technique is patterned after the Informal Reading Inventory devised by Betts (1946). Like the I.R.I., this new technique measures word pronunciation and comprehension. Betts' method uses materials of known difficulty to be read by readers of unknown reading ability in order to select a book for a given reader. In contrast, the Cloze Informal Reading Inventory uses materials of unknown difficulty to be read by readers of known reading ability in order to evaluate the readability of a book for a given indivdual or class. Also, the I.C.R.I. uses the cloze procedure to obtain a more valid measure of comprehension than can be obtained by teacher made test questions used in the I.R.I.

Subjects: Six average readers were chosen from classes at various grade levels from four through six in a public school system. At each grade level, three boys and three girls were selected who had Paragraph Meaning sub-test scores on the Standford Reading Achievement Test between the 40th and the 60th percentile ranks for their respective grades.

Materials: Three 1,000 word narratives were chosen from three different books for children. A 50 item cloze test was constructed (as previously described) based on a 250 word segment of each passage. The first 150 running words of the 250 word segment was used as an oral reading test. Two comprehension questions were devised for each oral reading test. These questions were not used to measure comprehension, but to establish the proper set for purposeful oral reading.

Data Gathering Procedures: Selected cloze tests and oral reading tests covering the same passages were administered to groups of six pupils at different grade levels.[2] Each group of six subjects first took a cloze test and then, one week later, each individual pupil read the 150 word passage orally to the teacher and answered the comprehension questions. The time between cloze and oral reading tests over the same passage was used to allow for forgetting. Fry and Dale-Chall readabilities were determined for each article.

Analysis: Scoring of cloze tests was based on the replacement of the exact word which had been deleted. On the oral reading tests, mispronunciations, substitutions, omissions, pauses, and insertions were counted as mistakes. Mean percentage scores were found for each group on both cloze and oral reading tests. The criterion for instructional level was a 44% cloze score and a 95% oral reading score.

Results: Table II shows the mean scores of each group of six average readers who took the test on a particular selection.

Table II

Mean Percentage Scores for the Informal
Cloze Readability Inventory

Selections	Cloze Inventory	Student Grade Level		
		4th	5th	6th
Danger Beneath the Ice.....	Comprehension %	11		53
	Oral Reading %	90		98
Goliath II.....	Comprehension %	13	53	59
	Oral Reading %	84	98	95
Penny and the White Horse....	Comprehension %			48
	Oral Reading %			98

[2]Tests were not administered at each grade level due to problems of time and convenience. This report is based on a larger study involving other subjects at various grade levels.

Visual inspection of the mean percentage scores given in Table II shows that a six-grade instructional level was found for "Danger Beneath the Ice" and "Penny and the White Horse." A fifth grade instructional level was found for "Goliath II." The progression of cloze percentages for "Goliath II" from fourth through sixth grades attests to the validity of this technique even though results are based upon a very small sample of six subjects at each grade level. It is interesting to note that this article meets the comprehension criterion for independent reading at the sixth grade level, but fails to pass the oral reading criterion. This illustrates one of the values of this technique over the Whole Class Procedure.

A comparison of the instructional level as determined by the Informal Cloze Readability Inventory with the readability grade levels measured by the Dale-Chall and Fry formulae is presented in Table III.

Table III

Comparison of Three Readability Assessments

Selections	Grade Level		
	Cloze Instructional	Fry Formula	Dale-Chall Formula
Goliath II	5	5	4 and below
Danger Beneath the Ice	6	4	5-6
Penny and the White Horse	6	6	4 and below

For these subjects, cloze instructional level was not greatly different from the readability levels as measured by the two formulae. Among these two comparisons, there appears to be a slight tendency for cloze instructional level

to place the passages at a slightly higher grade level than the level obtained by the formulae.

In general, it may be concluded that valid readability assessments can be made using the Informal Cloze Readability Inventory based on a small number of subjects. This technique could be modified to select "average readers" based on local norms rather than national norms. If this were done, cloze readability measurements would be more valid in many school systems than estimates or the results of readability formulae. Other percentages related to the conventional independent level or newly established empirically determined readability levels can be used with the Informal Cloze Readability Inventory.

Summary and Conclusions

There are several problems related to this investigation which should be noted. First, unless a good, recently normed reading test is used, the results of these methods will be misleading. Second, more work is needed on the problem of cloze sampling. These studies used only one sample of 250 words for each article. Obviously, several smaller samples would be necessary and practical in evaluating an entire book. Third, more work is needed on the interpretation of cloze percentages on materials whose difficulty varies greatly from the reading levels of the readers. Fourth, more research is needed using a larger number of subjects. Finally, it is clear that better readability assessment awaits a definitive empirical determination of both instructional and independent reading level criteria.

In the writer's opinion, the cloze procedure is an excellent technique for assessing readability. In terms of validity, there is good reason for using it as a standard in relation to which other types of readability measurements may be evaluated. The use of cloze criterion reference scores in relation

to measured reading grade level greatly enhances the utility of cloze test readability assessments. In the last analysis, readable writing is desirable and important for the reader's sake. If a passage is not readable to an intended reader, it is not readable--no matter how good a formula score it may receive. The reader must be the judge, and the cloze procedure allows the reader to have the last word.

List of References

Bormuth, J. R. Cloze tests as measures of readability. Unpublished Ed. D. dissertation, Indiana University, 1962.

Bormuth, J. R. Comparable cloze and multiple-choice comprehension test scores. Journal of Reading, 1967, 10, 291-299.

Bormuth, J. R. Cloze test readability: criterion reference scores. Journal of Educational Measurement, 1968, 5, 189-196.

Bormuth, J. R. Empirical determination of the instructional reading level. International Reading Association Conference Proceedings, 1969a, 716-721.

Bormuth, J. R. Development of readability analysis. (United States Office of Education, Project No. 7-0052) Chicago: University of Chicago, 1969b.

Gallant, R. An investigation of the use of cloze tests as a measure of readability of materials for the primary grades. Unpublished doctoral dissertation, Indiana University, 1964.

Rankin, E. F. and Culhane, J. W. Comparable cloze and multiple-choice comprehension test scores. Journal of Reading, 1969, 13, 193-198.

Rankin, E. F. and Culhane, J. W. One picture equals 1,000 words? Reading Improvement, 1970, 7, 37-40.

Sukeyori, S. A study of readability measurement-application of cloze procedure to Japanese Language. (English Abstract) Japanese Journal of Psychology, 1957, 28, 67-73.

Taylor, W. L. "Cloze" procedure: a new tool for measuring readability. Journalism Quarterly, 1953, 30, 415-438.